Tasmanian School of Business and Economics
University of Tasmania

Discussion Paper Series N 2018-04

An Unobserved Component Modeling Approach to Evaluate Multi-horizon Forecasts

Jing Tian

University of Tasmania, Australia

Thomas Goodwin

TasFood Ltd., Australia

ISBN 978-1-925646-49-8

An Unobserved Component Modeling Approach to Evaluate Multi-horizon Forecasts

Jing Tian *

University of Tasmania, CAMA

Thomas Goodwin

TasFood Ltd.

This version: September 2018

Abstract

We propose an unobserved modeling framework to evaluate a set of forecasts that target the same variable but are updated along the forecast horizon. The approach decomposes forecast errors into three distinct horizon-specific processes, namely, bias, rational error and implicit error, and attributes forecast revisions to corrections for these forecast errors. By evaluating multi-horizon daily maximum temperature forecasts for Melbourne, Australia, we demonstrate how this modeling framework can be used to analyze the dynamics of the forecast revision structure across horizons. Understanding forecast revisions is critical for weather forecast users to determine the optimal timing for their planning decisions.

JEL classification: C32; C53

Keywords: Decision making, decomposition, evaluating forecasts, state space models, weather forecasting

*Corresponding author: Jing Tian, Tasmanian School of Business and Economics, Private Bag 84, University of Tasmania, Hobart TAS 7001, Australia. Tel: +61 3 62262323, fax: +61 3 62267587. Email: jing.tian@utas.edu.au. We thank Dean Croushore, Jan Jacobs, Mardi Dungey and anonymous referees for detailed comments on the previous versions of the paper, and the seminar participants at University of Melbourne, Monash University, University of Adelaide, University of Queensland and Queensland University of Technology for helpful discussions.

1 Introduction

Forecasts for the same target are often provided multiple times before the target is realized. One of the most familiar examples relevant to everyday life is weather forecasts. The maximum and minimum temperatures and precipitation for a given day are forecast a few days in advance, and the public receives updates of these forecasts at least once each day before the target day. These multi-horizon weather forecasts play a significant role in decision-making processes related to a wide range of economic activities (see Dell et al. (2014)). For instance, agricultural practices, such as irrigation schedules and timing for harvest operations, rely on updated weather forecasts. Electricity generators receive electricity demand forecasts multiple times before dispatches, and these demand forecasts are often updated as weather forecasts are updated. To maximize the contributions of multi-horizon weather forecasts to planning decisions, it is critical that forecast users understand the forecast revision structure across horizons. Wang & Cai (2009) show that incorporating 7-day weather forecasts can increase crop net profits by 20%, but that "perfect" two-week weather forecasts in the form of actual weather data can achieve a 42% profit increase. Their study acknowledges the economic significance of using longer-horizon weather forecasts in planning but also suggests that the consequent economic benefits may depend on the trade-offs between the length of the forecast horizon and the amount of relevant information contained in long-horizon forecasts. Therefore, knowledge of whether forecast revisions contain information and how that information evolves in a sequence of forecast revisions is crucial for forecast users to determine the optimal timing for their decision making.

This paper provides a framework for forecast users, who may have little knowledge about how forecasts are generated, to improve their understanding of the structure of forecast errors and revisions across forecast horizons. Our modeling approach describes a general form of multi-horizon forecast errors with three distinct horizon-specific processes: rational forecast

error based on Muth's (1961) rational expectation hypothesis, implicit forecast error based on Mill's (1957) implicit expectation hypothesis, and systematic forecast bias. Under this general forecast error structure, we show that forecast revisions along horizons involve reducing the rational forecast error by incorporating newly available information, adjusting for implicit forecast error that is uncorrelated with the target, and correcting for systematic bias.

In the literature, the quality of multi-horizon forecasts is often assessed through rationality tests. Nordhaus (1987) introduce the concept of weak forecast efficiency to evaluate whether forecasts for the same target are rational. Weak-form efficiency requires that multi-horizon forecasts undergo revisions that are independent of past revisions and past forecast errors. A number of testing approaches have been built on this definition. Clements (1997) considers the scenario where only a small number of fixed-event forecasts are available and proposes pooling series of multi-horizon forecast across multiple target variables to conduct more powerful tests of weak-form efficiency. Clements & Taylor (2001) extend this approach to allow for non-normally distributed forecast revisions. Davies & Lahiri (1995, 1999) and Davies et al. (2011) focus on a three-dimensional panel data approach in which multi-horizon forecasts are produced by multiple forecasters and develop tests of rationality in a generalized method of moments framework.

Deviating from the commonly used testing approach for rationality, this paper proposes a modeling framework for evaluating a sequence of revised forecasts of the same target. We address the benefits of utilizing our modeling approach from the following aspects.

First, because the specifications of forecasts subject to a single type of error are nested by the specifications of multiple sources of error, model selection methods, such as information criteria and log likelihood tests, can be employed to identify the best fitted error structure of a given set of multi-horizon forecasts. When rational forecast error is estimated to be effectively the sole error type, the multi-horizon forecasts are seen to be rational since revisions are made purely for the purpose of adopting newly available information. Therefore, our

modeling approach provides an alternative method to rationality tests to evaluate the overall performance of multi-horizon forecasts.

Second, the estimation of our unobserved component model that explicitly specifies an error structure provides the estimated magnitude of each source of forecast revision at all horizons. Exploring the revision structure across horizons reveals the dynamics of how information contained in the forecast revisions evolves when approaching the target time. Forecast users are then able to identify the timing of the arrival of the largest amount of new information, which may help them to evaluate the trade-offs between early planning and information in long-horizon forecasts and to choose the optimal horizon forecast for making their planning decisions. The importance of understanding the changes in the information contents of forecasts over horizons has also been addressed by Isiklar & Lahiri (2007).

We demonstrate our model-based multi-horizon forecast evaluation approach using a real-time dataset of daily maximum temperature forecasts for Melbourne, Australia. Our results suggest that the weather forecasts revised at a daily frequency up to 14 days before the target day contain both rational and implicit errors. The composition of the sources for forecast revisions changes along the forecast horizon, and the incorporation of newly available information becomes the dominant attribute to revisions within a 7-day horizon.

We show the usefulness of our model-based evaluation approach by illustrating the value of the major upgrade to the National Center for Environmental Prediction's (NPEC) Numerical Weather Prediction (NWP) model on May 22, 2012. Subsample estimation results show that the upgrade provides more relevant information in the long-horizon forecasts and shifts the timing of the largest information arrival one day earlier, from 6 days out before the upgrade to 7 days out after the upgrade.

The remainder of this paper is structured as follows. In Section 2, we propose a model of multi-horizon forecasts that contains multiple sources of forecast errors. In Section 3, we cast our models of various error structures in a state space form. Section 4 conducts simulations.

We evaluate multi-horizon weather forecasts as an empirical illustration of our approach in Section 5, and Section 6 concludes this paper.

2 Multifarious errors in multi-horizon forecasts

In this section, we develop a model of multi-horizon forecasts that contain multiple sources of error. We begin by outlining our assumptions and describing our most general model specification. In the subsections that follow, we explain how the target variable and three different types of forecast errors can be modeled as unobserved components.

2.1 A model of multi-horizon forecasts

Suppose that forecast users are given a sequence of numerical multi-horizon forecasts for the same target event $\tilde{y}_t$. Let the forecasts made at h periods earlier, i.e., at time $t-h$, be denoted by $\hat{y}_{t|t-h}$. The longest-horizon forecasts are provided at horizon $h = H$, where $H > 1$. As the target event approaches, forecast users observe a number of forecast revisions between two adjacent updating points, $d_{t|h-1,h} = \hat{y}_{t|t-(h-1)} - \hat{y}_{t|t-h}$, for the same target $\tilde{y}_t$.

The target $\tilde{y}_t$ is assumed to be a covariance-stationary process, which can be generalized to an MA form of white noise according to the Wold decomposition, that is

$$\begin{aligned} &\tilde{y}_t = \mu_{\tilde{y}} + \textstyle\sum_{i=0}^{\infty} \theta_i u_{t-i}, \\ &\text{where } \theta_0 = 1, \textstyle\sum_{i=1}^{\infty} \theta_i^2 < \infty \text{ and } u_t \sim i.i.d.(0, \sigma_u^2). \end{aligned} \tag{1}$$

This moving average representation describes the data generating process as an information accumulation process, and hence, is often used for the dynamics of the target event in the fixed-event forecasting literature (see for example, Muth (1985), Isiklar & Lahiri (2007), and Chang et al. (2013)). We re-parameterize the linear combination of white noise in equation

(1) as

$$\sum_{i=0}^{\infty} \theta_i u_{t-i} = \sum_{i=0}^{\infty} \omega_{i,t} = \sum_{i=0}^{\infty} \sigma_{\omega_i} \eta_{\omega_{i,t}}, \tag{2}$$

where $\eta_{\omega_{i,t}} \sim N(0,1)$ and are uncorrelated across time and horizons. Therefore, $\sum_{i=0}^{\infty} \omega_{i,t}$ represents the cumulative effect of all past unanticipated shocks in the realization of $\tilde{y}_t$, and at each of the i periods away from the target time t, the effect of the shock $\omega_{i,t} \sim i.i.d.N(0, \sigma^2_{\omega_i})$. Combining equations (1) and (2), the target event $\tilde{y}_t$ with a finite sample can be expressed as

$$\tilde{y}_t = \mu_{\tilde{y}} + \sum_{i=0}^{t-1} \omega_{i,t}. \tag{3}$$

After subtracting the unconditional mean $\mu_{\tilde{y}}$ from $\tilde{y}_t$, we can separate the information accumulation $\sum_{i=0}^{t-1} \omega_{i,t}$ in the above equation into two sections,

$$\tilde{y}_t - \mu_{\tilde{y}} = \sum_{i=H}^{t-1} \omega_{i,t} + \sum_{i=0}^{H-1} \omega_{i,t}, \tag{4}$$

where the first term on the right-hand side of the equation includes unanticipated information that occur from the first observation up to the time when the longest horizon forecast is made at the H period before the target time t, and the second term includes information occurring subsequently until the realization of the target.

Our aim is to characterize a sequence of multi-horizon forecasts $\left(\hat{y}_{t|t-H}, \hat{y}_{t|t-H+1}, ..., \hat{y}_{t|t-1}\right)'$ for the same target $\tilde{y}_t$ with a model that specifies one or more sources of forecast error. We decompose forecasts $\hat{y}_{t|t-h}$ as

$$\hat{y}_{t|t-h} = \tilde{y}_t + \beta_h + \nu_{t|t-h} + \zeta_{t|t-h}, \tag{5}$$

or a "demeaned" version for multi-horizon forecasts is

$$\hat{y}_{t|t-h} - \mu_{\tilde{y}} = \tilde{y}_t - \mu_{\tilde{y}} + \beta_h + \nu_{t|t-h} + \zeta_{t|t-h}, \tag{6}$$

where $\tilde{y}_t - \mu_{\tilde{y}}$ is defined in equation (4).

Equation (5) or (6) indicates that the forecasting errors $\tilde{y}_t - \hat{y}_{t|t-h}$ nest three types of error: horizon specific bias, $-\beta_h$; accumulated unanticipated shocks during the forecast horizon after the forecasts are made, $-\nu_{t|t-h}$; and the error component unrelated to the target, $-\zeta_{t|t-h}$. We discuss each type of error in the following subsections.

2.2 Rational forecast errors

The rational expectations hypothesis of Muth (1961) suggests that rational forecasters form their expectations by effectively using all of the available information; hence, the only source of error is unanticipated information received after the forecasts are made. At an h horizon prior to the target time t, a rational forecaster with no access to unanticipated information that occurs over the forecast horizon will produce a rational forecast:

$$\hat{y}_{t|t-h} = \tilde{y}_t - \sum_{i=0}^{h-1} \omega_{i,t} = \tilde{y}_t + \nu_{t|t-h}. \tag{7}$$

The rational forecast error is the accumulation of $\omega_{i,t}$ over the horizon, that is, $-\nu_{t|t-h}$.

We now analyze the properties of the rational forecast error. As defined in Section 2.1, the effect of each unanticipated shock $\omega_{i,t} = \sigma_{\omega_i}\eta_{\omega_i,t}$, where $\eta_{\omega_i,t} \sim i.i.d.N(0,1)$ and $\sigma_{\omega,i} \geq 0$. In the case of $\sigma_{\omega_i} = 0$, there is no unanticipated shock occurring at horizon i prior to the target time t. The rational forecast error $\nu_{t|t-h}$ has zero expectation and possesses a number of distinct properties. First, equation (3) indicates that rational forecast error is correlated with the target value, i.e., $cov(\tilde{y}_t, \nu_{t|t-h}) \neq 0$. Second, the unanticipated shocks that occur after the forecasts are made are future information yet to be incorporated into the rational forecasts (as seen in equation (7), in which the future information is deducted from the target value); therefore, $cov(\hat{y}_{t|t-h}, \nu_{t|t-h}) = 0$. Moreover, the variance of the rational forecast error, computed as $\mathrm{E}(\nu^2_{t|t-h}) = \sum_{i=0}^{h-1} \sigma^2_{\omega_i,t}$, is non-decreasing as the forecast horizon h increases.

This last property is intuitively appealing, as we expect there to be less relevant information available to forecasters at longer horizons. A decline in relevant information will cause an increase in forecast uncertainty and an associated increase in the variation of rational forecast errors. Finally, on the basis of the first two properties, the variance of the target is no less than the variance of the rational forecast.

2.3 Implicit forecast error

Forecasts may deviate from rational forecasts. Mills (1957) introduces the concept of "implicit" expectation to describe the motivation of firms to hold inventories that deviate from rational expectations of demand. This hypothesis was empirically tested by Lovell (1986). In the scenario where rational forecasts of future demand are declining, a firm may still face a relatively high demand forecast and hold inventories in order to reduce fluctuations in production. Such forecast errors are formed due to consideration of the high costs associated with rapid changes in production. They are uncorrelated with the actual demand but correlated with the demand forecasts (as they are introduced by forecasters). This covariance property of implicit forecast error is opposite from that possessed by rational forecast error.

Considering implicit forecast error as noise introduced into forecasts, the level of noise may be different across forecast horizons. It is modeled by

$$\zeta_{t|t-h} = \sigma_{\zeta_h} \eta_{\zeta_h,t}, \tag{8}$$

where $\eta_{\zeta_h,t} \sim i.i.d.\ \mathrm{N}(0,1)$. The following covariance assumptions enable us to differentiate them from the rational forecast error: 1) $\mathrm{Cov}(\hat{y}_{t|t-h}, \zeta_{t|t-h}) \neq 0$ and 2) $\mathrm{cov}(\tilde{y}_t, \zeta_{t|t-h}) = 0$. If multi-horizon forecasts are only subject to implicit forecast error, the variance of the target must be less than the variance of the implicit forecasts. In contrast to the monotonic pattern of rational forecast error, the variance of implicit forecast error, $\mathrm{E}(\zeta^2_{t|t-h}) = \sigma^2_{\zeta_h}$, may either

increase or decrease as the forecast horizon shortens.

Forecasters may introduce implicit error into their forecasts unintentionally. For instance, information used by forecasters may contain measurement errors unrelated to the target, preventing forecasters from using the information efficiently (Lovell, 1986). As pointed out by Isiklar & Lahiri (2007), the measurement error may be driven by noises in the data generating process, delays in data releases and data revisions.

2.4 Bias

Both rational and implicit forecast errors have zero expectations; however, a large body of research has empirically found non-zero bias in a wide range of forecasts, such as financial analysts' earning forecasts, gross debt forecasts and inflation forecasts. It is now well accepted that forecasts can remain rational in the presence of bias if the forecasters' loss function is asymmetric (see Christoffersen & Diebold (1997), Lim (2001) and Patton & Timmermann (2007)).

Therefore, we allow for a non-zero systematic bias as the third type of forecast error. We assume that the bias β_h is time-invariant but horizon specific[1] so that empirically, bias can easily be distinguished from rational and implicit forecast errors.

3 A state space representation of multi-horizon forecasts

3.1 A general structure

In this section, we cast the unobserved component model for multi-horizon forecasts that contain multiple sources of error in a state space form. The time-invariant state space model

[1]Note that, although the forecasters' learning process may indicate time-varying bias, in this paper, we restrict the bias to being constant to avoid any potential identification issues. The time-invariant feature of bias is also consistent with Davies & Lahiri (1995).

consists of a set of measurement equations and a set of transition equations, that is,

$$\boldsymbol{y}_t = \boldsymbol{Z}\boldsymbol{\alpha}_t \tag{9}$$

$$\boldsymbol{\alpha}_t = \boldsymbol{T}\boldsymbol{\alpha}_{t-1} + \boldsymbol{R}\boldsymbol{\eta}_t. \tag{10}$$

The specifications in equation (9) and (10) build on the multifarious errors model in the form of equation (6) and the definitions and covariance properties of different types of errors discussed in Section 2. Considering a case in which the observed value of the target is the same as the target value, i.e., $y_t = \tilde{y}_t$, we have equal unconditional means of the observed and the target value, i.e., $\mu_y = \mu_{\tilde{y}}$.

We first estimate $\mu_{\tilde{y}}$ by calculating the sample mean of the observed value of the target event and denote it with $\hat{\mu}_y$. Subtracting $\hat{\mu}_y$ from the multi-horizon forecasts and the observed targeted values, we have the measurement vector, $\boldsymbol{y}_t = [\hat{y}_{t|t-H} - \hat{\mu}_y, \hat{y}_{t|t-(H-1)} - \hat{\mu}_y, \ldots, y_t - \hat{\mu}_y]'$.[2] In our most general model that specifies all three types of forecast errors, the state vector is partitioned as follows

$$\boldsymbol{\alpha}_t = \begin{bmatrix} \tilde{y}_t^* & \boldsymbol{\beta_h}' & \boldsymbol{\nu}'_{t|t-h} & \boldsymbol{\zeta}'_{t|t-h} \end{bmatrix}', \tag{11}$$

where $\tilde{y}_t^*$ denotes the demeaned values of the target variable, and $\boldsymbol{\alpha}_t$ has length $(1+H+H+H)$.

[2] Alternatively, one can establish the measurement equation based on equation (5) without the need to subtract μ_y from the forecasts and the observed targeted values. In this case, the right-hand side of the state equation for $\tilde{y}_t$ has an intercept $\mu_{\tilde{y}}$ that requires estimation. The GAUSS software allows for direct estimation of the intercept in state equations, but we use the Econometrics Toolbox from MATLAB to build and estimate the state space model, which strictly follows the form of equation (10). Although we could nominate $\mu_{\tilde{y}}$ as one of the unobserved state variables, its time-invariant property is then identical to $\boldsymbol{\beta_h}$, resulting in an identification problem in the state space model estimation.

The associated measurement equation is

$$\boldsymbol{y}_t = \begin{bmatrix} \boldsymbol{Z}_1 & \boldsymbol{Z}_2 & \boldsymbol{Z}_3 & \boldsymbol{Z}_4 \end{bmatrix} \begin{bmatrix} \tilde{y}_t^* \\ \boldsymbol{\beta_h} \\ \boldsymbol{\nu}_{t|t-h} \\ \boldsymbol{\zeta}_{t|t-h} \end{bmatrix}, \tag{12}$$

where $\boldsymbol{Z} = [\boldsymbol{Z}_1 \quad \boldsymbol{Z}_2 \quad \boldsymbol{Z}_3 \quad \boldsymbol{Z}_4]$ is a partitioned matrix conforming to the unobserved components of the state vector; $\boldsymbol{Z}_1 = \mathbf{1}_{(H+1)}$ (which is an $(H+1) \times 1$ vector of ones) is related to the target variable component; and $\boldsymbol{Z}_2$, $\boldsymbol{Z}_3$ and $\boldsymbol{Z}_4$ are each $[\boldsymbol{I}_H, \mathbf{0}'_{1\times H}]'$ (an $H \times H$ identity matrix atop a conformably defined vector of zeros) and are related to the bias component, the rational error component and the implicit error component, respectively. This measurement equation reflects the forecast decomposition of equation (6) and also equates the (demeaned) observed value of the target variable y_t with the (demeaned) "true" unobserved target value $\tilde{y}_t$.

Transition equations describe the dynamics of the unobserved components in terms of the state vector

$$\begin{bmatrix} \tilde{y}_t^* \\ \boldsymbol{\beta}_h \\ \boldsymbol{\nu}_{t|t-h} \\ \boldsymbol{\zeta}_{t|t-h} \end{bmatrix} = \begin{bmatrix} T_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{T}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{T}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{T}_4 \end{bmatrix} \begin{bmatrix} \tilde{y}_{t-1}^* \\ \boldsymbol{\beta}_h \\ \boldsymbol{\nu}_{t-1|t-1-h} \\ \boldsymbol{\zeta}_{t-1|t-1-h} \end{bmatrix}$$
$$+ \begin{bmatrix} R_1 & \mathbf{0} & \boldsymbol{R}_3 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{R}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\boldsymbol{U} \cdot \mathrm{diag}(\boldsymbol{R}_3) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathrm{diag}(\boldsymbol{R}_4) \end{bmatrix} \begin{bmatrix} \eta_{\xi,t} \\ \boldsymbol{\eta}_{\beta_h,t} \\ \boldsymbol{\eta}_{\omega_h,t} \\ \boldsymbol{\eta}_{\zeta_h,t} \end{bmatrix}, \tag{13}$$

where $\boldsymbol{U}$ is an $H \times H$ matrix with zeros below the main diagonal and each of the remaining elements equal to one, and $[\eta_{\xi,t}, \boldsymbol{\eta}'_{\beta_h,t}, \boldsymbol{\eta}'_{\omega_h,t}, \boldsymbol{\eta}'_{\zeta_h,t}]' \sim i.i.d.N(\mathbf{0}, \boldsymbol{I})$. Looking at the partitions of the transition coefficient $\boldsymbol{T}$, T_1 is a scalar; and $\boldsymbol{T}_2$, $\boldsymbol{T}_3$ and $\boldsymbol{T}_4$ are all $H \times H$. $\mathbf{0}$ is an $H \times H$ null matrix. The partitions of the loading coefficients $\boldsymbol{R}$ have similar dimensions to those in $\boldsymbol{T}$, and $\boldsymbol{R}_3$ and $\boldsymbol{R}_4$ are all $1 \times H$.

3.2 Specifications for the target variable

To discuss the block for the target $\tilde{y}_t$, we start from equation (4) in section 2.2,

$$\tilde{y}_t^* = \sum_{i=H}^{t-1} \omega_{i,t} + \sum_{i=0}^{H-1} \omega_{i,t}.$$

All of the unanticipated new information $\omega_{i,t}$ in the second summation of the above equation occurs after the longest horizon forecasts are made and can be identified by the observed set of multi-horizon forecasts. However, the information included in the first summation occurs during the period spanning from the time when the longest horizon forecast was made all the way back to the first time point, that is, $\omega_{i,t}$ for $i = H, H+1, ..., t-1$. This information is a part of the rational forecasts at every given forecast horizon and cannot be separately identified from the summation. Therefore, we let $\sum_{i=H}^{t-1} \omega_{i,t} = \sigma_\xi \eta_{\xi,t}$ where $\eta_{\xi,t} \sim i.i.d.N(0,1)$.

The block for the demeaned target variable in equation (13) is hence given as

$$\tilde{y}_t^* = \begin{bmatrix} \sigma_\xi & \sigma_{\omega_{H-1}} & \sigma_{\omega_{H-2}} & \dots & \sigma_{\omega_1} & \sigma_{\omega_0} \end{bmatrix} \begin{bmatrix} \eta_{\xi,t} \\ \boldsymbol{\eta}_{\omega_h,t} \end{bmatrix}. \tag{14}$$

Specifically, the transition coefficients $T_1 = 0$, the loading coefficients $R_1 = \sigma_\xi$ and $\boldsymbol{R}_3$ is a $1 \times H$ vector of $[\sigma_{\omega_{H-1}}, \sigma_{\omega_{H-2}}, \dots, \sigma_{\omega_1}, \sigma_{\omega_0}]$.

3.3 Specifications for forecasting errors

We now explain the blocks for the unobserved forecast errors. Since the bias is a horizon-specific constant that does not vary over time, we set $\boldsymbol{T}_2$ as an $H \times H$ identity matrix. On the basis of the discussion in section 2.2, at each horizon $\nu_{t|t-h} = -\sum_{i=0}^{h-1} \omega_{i,t} = -\sum_{i=0}^{h-1} \sigma_{\omega_i} \eta_{\omega_i,t}$; therefore, we let $\boldsymbol{T}_3$ be an $H \times H$ null matrix and $\boldsymbol{R}_3 = [\sigma_{\omega_{H-1}}, \sigma_{\omega_{H-2}}, \ldots, \sigma_{\omega_1}, \sigma_{\omega_0}]$. The rational forecast error component is then given by

$$\boldsymbol{\nu}_{t|t-h} = -\boldsymbol{U} \cdot \text{diag}(\boldsymbol{R}_3) \cdot \boldsymbol{\eta}_{\omega_h,t} = \begin{bmatrix} -\sigma_{\omega_{H-1}} & -\sigma_{\omega_{H-2}} & \ldots & -\sigma_{\omega_1} & -\sigma_{\omega_0} \\ 0 & -\sigma_{\omega_{H-2}} & \ldots & -\sigma_{\omega_1} & -\sigma_{\omega_0} \\ \vdots & \ddots & \ddots & -\sigma_{\omega_1} & -\sigma_{\omega_0} \\ 0 & \ldots & \ldots & 0 & -\sigma_{\omega_0} \end{bmatrix} \begin{bmatrix} \eta_{\omega_{H-1},t} \\ \eta_{\omega_{H-2},t} \\ \vdots \\ \eta_{\omega_1,t} \\ \eta_{\omega_0,t} \end{bmatrix}. \quad (15)$$

Jacobs & Van Norden (2011) use this restricted specification of the loading matrix to model the rational data revision process.

The implicit forecast error is horizon specific and uncorrelated across horizons. Thus, the matrix $\boldsymbol{T}_4$ is a null matrix, and the dynamics of the implicit component are completely described by $\text{diag}(\boldsymbol{R}_4)$, where $\boldsymbol{R}_4$ is a row vector of standard deviations of the implicit forecast error, i.e., $[\sigma_{\zeta_H}, \sigma_{\zeta_{H-1}}, ..., \sigma_{\zeta_1}]$. The equation below describes the process of implicit forecast error:

$$\boldsymbol{\zeta_{t|t-h}} = \text{diag}(\boldsymbol{R}_4) \cdot \boldsymbol{\eta}_{\zeta,t} = \begin{bmatrix} \sigma_{\zeta_H} & 0 & \ldots & 0 \\ 0 & \sigma_{\zeta_{H-1}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \sigma_{\zeta_1} \end{bmatrix} \begin{bmatrix} \eta_{\zeta_H,t} \\ \eta_{\zeta_{H-1},t} \\ \vdots \\ \eta_{\zeta_1,t} \end{bmatrix}. \quad (16)$$

The most general model represented by equations (12) and (13) nests several simpler

forecast error structures. For example, by removing the blocks for bias $\boldsymbol{\beta}_h$, we model multi-horizon forecasts to contain only rational and implicit errors. The transition equation is

$$\begin{bmatrix} \tilde{y}_t \\ \boldsymbol{\nu}_{t|t-h} \\ \boldsymbol{\zeta}_{t|t-h} \end{bmatrix} = \begin{bmatrix} T_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{T}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{T}_4 \end{bmatrix} \begin{bmatrix} \tilde{y}_{t-1} \\ \boldsymbol{\nu}_{t-1|t-1-h} \\ \boldsymbol{\zeta}_{t-1|t-1-h} \end{bmatrix} + \begin{bmatrix} R_1 & \boldsymbol{R}_3 & \mathbf{0} \\ \mathbf{0} & -\boldsymbol{U} \cdot \text{diag}(\boldsymbol{R}_3) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \text{diag}(\boldsymbol{R}_4) \end{bmatrix} \begin{bmatrix} \eta_{\xi,t} \\ \boldsymbol{\eta}_{\omega_h,t} \\ \boldsymbol{\eta}_{\zeta_h,t} \end{bmatrix}. \quad (17)$$

A state space form for multi-horizon forecasts subject purely to rational forecast error can then be obtained by further restricting $\boldsymbol{Z}_4$ and $\boldsymbol{R}_4$ to be null matrices, and the transition equation is,

$$\begin{bmatrix} \tilde{y}_t \\ \boldsymbol{\nu}_{t|t-h} \end{bmatrix} = \begin{bmatrix} T_1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{T}_3 \end{bmatrix} \begin{bmatrix} \tilde{y}_{t-1} \\ \boldsymbol{\nu}_{t-1|t-1-h} \end{bmatrix} + \begin{bmatrix} R_1 & \boldsymbol{R}_3 \\ \mathbf{0} & -\boldsymbol{U} \cdot \text{diag}(\boldsymbol{R}_3) \end{bmatrix} \begin{bmatrix} \eta_{\xi,t} \\ \boldsymbol{\eta}_{\omega_h,t} \end{bmatrix}. \quad (18)$$

Alternatively, in the absence of unanticipated new information over the forecast horizon, revised multi-horizon forecasts only consist of forecast error uncorrelated with the target. In this case, we set $\boldsymbol{Z}_3$ to be a null matrix, and let $\boldsymbol{R}_3$ be a $1 \times H$ vector of zeros. The transition equation for modeling the pure implicit forecasts is then given by

$$\begin{bmatrix} \tilde{y}_t \\ \boldsymbol{\zeta}_{t|t-h} \end{bmatrix} = \begin{bmatrix} T_1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{T}_4 \end{bmatrix} \begin{bmatrix} \tilde{y}_{t-1} \\ \boldsymbol{\zeta}_{t-1|t-1-h} \end{bmatrix} + \begin{bmatrix} R_1 & \mathbf{0} \\ \mathbf{0} & \text{diag}(\boldsymbol{R}_4) \end{bmatrix} \begin{bmatrix} \eta_{\xi,t} \\ \boldsymbol{\eta}_{\zeta_h,t} \end{bmatrix}. \quad (19)$$

3.4 Estimation and analysis of the forecast revision structure

We check the various state space representations of multi-horizon forecasts against the sufficient conditions of controllability and observability provided by Harvey (1989) and Jacobs & van Norden (2007). The parameters in the models discussed in the previous subsection are identified.

The maximum likelihood estimator with the Kalman filter is used to estimate the unobserved component models. Since the specifications of the multi-horizon error structure are nested, model selection methods, such as log likelihood ratio tests and the conventional information criteria, can be implemented to determine the best fitted error structure.

The estimation of multifarious error models helps to provide insight into the sources of forecast revisions across horizons. Suppose that the multi-horizon forecasts consist of all three types of forecast errors and are modeled by equations (12) and (13). The differences between $\hat{y}_{t|t-(h-1)}$ and $\hat{y}_{t|t-h}$, namely, the marginal revisions made between two updating points $t-(h-1)$ and $t-h$, are derived as follows,

$$\begin{aligned} d_{t|h-1,h} &= \hat{y}_{t-(h-1)} - \hat{y}_{t|t-h} \\ &= \beta_{h-1} - \beta_h + \sigma_{\omega_{h-1}} \eta_{\omega_{h-1},t} + \zeta_{t|t-(h-1)} - \zeta_{t|t-h}. \end{aligned} \tag{20}$$

Therefore, the mean squared forecast revisions (MSFR) can be decomposed as

$$\begin{aligned} \text{MSFR}_{t|h-1,h} &= E(d^2_{t|h-1,h}) \\ &= (\beta_{h-1} - \beta_h)^2 + \sigma^2_{\omega_{h-1}} + \sigma^2_{\zeta_{h-1}} + \sigma^2_{\zeta_h}. \end{aligned} \tag{21}$$

Equation (21) indicates that marginal revisions are made to correct bias (first term), to

adopt newly available information (second term) and to adjust for implicit errors that are uncorrelated with the target (the last two terms).

4 Simulation

In this section, we undertake a simulation exercise to evaluate the performance of the proposed modeling approach for identifying forecast error compositions across horizons. We let the observed target variable y_t follow a simple AR(1) process

$$y_t = \mu_y + \phi(y_{t-1} - \mu_y) + \epsilon_t, \tag{22}$$

and $\epsilon_t \sim i.i.d.N(0, \sigma_\epsilon^2)$, where $\sigma_\epsilon = (1-\phi^2)^{1/2}\sigma_y$. Calibrating the parameters to be close to the data of the daily maximum temperature in Melbourne, Australia, we set $\mu_y = 21$, $\phi = 0.75$ and $\sigma_y = 6$. The standard deviation of ϵ_t is hence $\sigma_\epsilon = 3.97$.

We generate forecasts made at four horizons $h = 4, 3, 2,$ and 1 to contain two types of forecast errors: rational and implicit errors. To do so, we first calculate rational forecasts $\mathrm{E}(y_{t|t-h})$ using equation (22) and then add implicit errors to the rational forecast at each horizon. The standard deviations of marginal information adoption by rational forecast at each horizon can be derived from iterating y_t in equation (22) as

$$y_t = \mu_y + \phi^h(y_{t-h} - \mu_y) + \phi^{h-1}\epsilon_{t-(h-1)} + \phi^{h-2}\epsilon_{t-(h-2)} + ... + \phi\epsilon_{t-1} + \epsilon_t. \tag{23}$$

The information relevant to y_t that is unavailable at horizon h but becomes observed at horizon $h-1$ is $\phi^{h-1}\epsilon_{t-(h-1)}$; hence, for $h = 4, 3, 2$ and 1, the standard deviations of marginal information adoption, $\sigma_{\omega_{h-1}}$, are $\sigma_{\omega_3} = 0.75^3 \times 3.97 = 1.67$, $\sigma_{\omega_2} = 0.75^2 \times 3.97 = 2.23$, $\sigma_{\omega_1} = 0.75 \times 3.97 = 2.98$, and $\sigma_{\omega_0} = 3.97$.

We consider two sets of values for the standard deviations of implicit error. First, we let

$\sigma_{\zeta_h} = \frac{h-1}{3}\sigma_y$ for $h = 4, 3, 2$ and 1. In this case, the magnitude of implicit error at the longest horizon $h = 4$ is as great as σ_y, and as the forecast horizon shrinks, implicit error declines. In the second data generating process, we consider a smaller value of $\sigma_{\zeta_h} = 3$ for all four horizons. We use a sample size of 300 observations and perform 1000 simulations.

In each of the simulations, we estimate three models: the most general three-error component model; the rational plus implicit error component model; and the rational error component model. Table 1, Figure 1 and Figure 2 report the simulation results of the estimated parameters in the rational plus implicit error component model, which is the correct model used in the data generating process. As indicated in Table 1, both the Akaike Information Criterion (AIC) and Baysian Information Criterion (BIC) choose the correct error structure with a successful rate of 99.7% for the first DGP and 98.3% for the second DGP. The average values of estimated parameters across 1000 simulations (which are reported in the second row of each panel in Table 1) are close to their true values (which are reported in the first row of each panel in Table 1).

[Insert Table 1]

Figures 1 and 2 provide the histograms of the simulation results for the first and the second DGP, respectively. In general, we observe bell-shaped distributions, and the estimated parameters are centered around their true values. In Figure 1, when there is no implicit error at the shortest horizon $h = 1$, the largest mass of $\hat{\sigma}_{\zeta_1}$ occurs at the true value of 0.

[Insert Figure 1]

[Insert Figure 2]

The simulation results suggest that this unobserved component modeling approach is able to identify the correct error structure and to provide reliable estimates for the magnitude of the information adoption and implicit forecast error across forecast horizons.

5 Evaluating Multi-horizon Weather Forecasts

In this section, we evaluate multi-horizon forecasts of the daily maximum temperature (degrees Celsius) for Melbourne, Australia to demonstrate how our model-based evaluation approach extracts the various types of forecast errors and forecast revision components and to suggest how these results could assist decision makers in planning.

Over the past few decades, meteorological services in Australia have produced increasingly accurate weather forecasts at ever-increasing forecast horizons (see Stern (2008) and Stern & Davidson (2015)). For decision makers who rely on weather conditions, the technological improvement of weather forecasts provides an opportunity to choose between using long-horizon weather forecasts and using short-horizon weather forecasts. We aim to provide insight into the revision process for daily maximum temperature forecasts over 14-day horizons so that decision makers can optimally time their planning decisions conditional on the weather forecast revision process.

5.1 Data

We retrieve the data from `http://www.weather-climate.com`, which consist of an experimental daily maximum temperature forecast series generated at multiple 14-day horizons and an observed maximum daily temperature series for Melbourne, Australia. The sample period runs from February 1, 2009 to December 31, 2014, comprising a total of 2159 days, with forecasts available at 14, 13,..., 2, and 1 day out from each observation date.

These experimental daily maximum temperature forecasts were produced in real-time using a forecast combination algorithm, as documented in Stern (2007) and Stern & Davidson (2015). A number of data sources are used to produce the combined forecasts, including the official forecasts from the Australian Bureau of Meteorology (BOM), the previous day's

maximum temperature forecasts, statistical forecasts[3], and climatological forecasts[4]. Table 2 describes the forecast combination weightings used to generate the daily maximum temperature forecasts. The combination weights vary depending on the length of the forecast horizon. For example, BOM does not publish forecasts of maximum temperature for more than 7 days out from the target date; hence heavier weights are imposed on the statistical and climatological forecasts that are based on long-run and large-scale weather forecasts.

[Insert Table 2]

Figure 3 plots the daily maximum temperature observations (light colored dotted line) and the corresponding meteorological forecasts generated 7 days (gray solid line) and 14 days (black solid line) out from the target observation date, spanning the full sample. This figure presents three main features. First, the observed maximum temperature series is more volatile than both forecast series, and the forecasts made 7 days before the target are more volatile than the forecasts made 14 days before the target. These observations meet Patton & Timmermann (2012)'s necessary conditions for rational forecasts, that is, the variance of short-horizon forecasts is no less than the variance of long-horizon forecasts and is bounded by the variance of the realization. Second, there is a permanent increase in the variations of the two forecast series beginning from mid-2012. Stern & Davidson (2015) discuss an improvement in forecast skill beginning from mid-2012, which they attribute to a major upgrade of the NWP models on May 22, 2012. In addition, the observed temperature series and the two forecast temperature series exhibit more variability on warmer days than on cooler days. Stern & Davidson (2015) note that the competing influence of warm dry winds from the Australian interior and cool moist winds from the Southern Ocean render temperature forecasting for

[3]Stern and Davidson (2015) provide a brief explanation of the statistical forecasts. These forecasts are computed for local weather based on the output of the long-range numerical weather prediction (NWP) models provided by the National Center for Environmental Prediction (NCEP) of the National Oceanic and Atmospheric Administration (NOAA). See Wilks (2011) for examples of statistical forecast methods for meteorological variables.

[4]Climatological forecasts are the averages of historical observations over many years.

Melbourne particularly challenging, and this influence is strongest during the warmer months.

[Insert Figure 3]

We report both full-sample and sub-sample means and the standard deviations of the maximum daily temperature observations and all fourteen horizon forecasts in Table 3[5]. The statistics confirm our observations regarding Figure 3.

[Insert Table 3]

5.2 Results

5.2.1 Forecast decomposition

We focus on three model specifications: the rational model that specifies unanticipated new information over the forecast horizon as the only source of forecast error; the rational plus implicit model that allows for some forecast errors to be unrelated to the target; and the bias+rational+implicit model that adds a horizon-specific but time-invariant systematic bias[6].

Table 4 reports the estimation results of the three alternative models. The top panel presents the estimated values of $\sigma_{\omega_{h-1}}$, which represents the marginal increase in information content owing to forecast revisions made at the shorter horizon $h-1$ when compared to the forecast of horizon h. The estimated standard deviations of the implicit errors at each horizon h, denoted by σ_{ζ_h}, are reported in the middle panel. These values capture the size of the noise uncorrelated with the targeted maximum daily temperature. The bottom panel shows the Kalman smoothed estimates of the horizon-specific forecast bias. We also report the log likelihood values, and the Akaike and Bayesian Information Criteria for each model.

[5]For the Southern hemisphere, we denote the period from September 21 to March 20 as warm days and the period from March 21 to September 20 as cool days.

[6]The estimation results of other alternative models, including the pure implicit model, the bias plus rational model, and the bias plus implicit model, can be provided upon request. The log likelihood values suggest that these three models are less preferred to the rational model, the rational plus implicit model and the bias+rational+implicit model reported here.

[Insert Table 4]

Since the rational model is nested by both multiple-error models, we can apply either log likelihood ratio tests or information criteria to identify the empirically preferred model(s). Both approaches reach the conclusion that multi-horizon forecasts of the maximum daily temperature for Melbourne up to 14 days out are subject to multiple types of forecast errors. Between the two multiple-error models, which have different state variables but the same number of unknown parameters, we can simply compare the maximum values of log likelihood. Since the rational and implicit model achieves a higher log likelihood value than that of the model with systematic bias, we prefer this bi-error structure and continue our further analysis based on the estimates of the rational plus implicit model [7].

The estimated value of $\sigma_{\omega_{h-1}}$ starts relatively low at 0.637°C in the revised forecast at the horizon of 13 days and then remains at a similar level until 9 days before the target date. It then increases in the subsequent revisions and peaks at the value of 1.560°C at the forecast horizon of 6 days. Subsequently, as the horizon shrinks further, the estimated $\sigma_{\omega_{h-1}}$ declines, suggesting decreasing marginal information content adoption within 5 days before the target day.

The estimated standard deviation of the implicit forecast error σ_{ζ_h} exhibits a different pattern. It is highest at the longest forecast horizon and gradually declines as the forecast time approaches 7 days before the target day. Revisions of the maximum temperature forecasts within a week of the target day lead to trivial implicit errors.

We use Figure 4 to depict the dynamics of the forecast revision structure across forecast horizons. With a modification of equation (21) for the rational plus implicit forecasts by

[7] Note that, for the bias+rational+implicit model, the estimated biases over all 14 horizons are significantly negative, but the estimates related to the rational revision and implicit forecast errors are virtually the same as those from the rational plus implicit model, except for the implicit forecast errors within 7 days before the target day. In both models, the estimated magnitudes of the implicit forecast errors within 7 days are significantly lower than those of the rational forecast errors. Therefore, our subsequent analysis, specifically the change in marginal information content in the revisions over forecast horizons, is not much affected by our modeling choice.

removing the bias term, we can decompose the value of $\text{MSFR}_{t|h-1,h}$ into $\sigma^2_{\omega_{h-1}}$ plus $\sigma^2_{\zeta_h}$ and $\sigma^2_{\zeta_{h-1}}$. Figure 4 shows how the sizes of these components for the revised daily maximum temperature forecasts evolve from 13 days to 1 day before the realization. The total length of each bar represents the size of $\text{MSFR}_{t|h-1,h}$. Within each bar, the black color represents the variation of newly adopted information in the updated forecast made at horizon $h-1$, and the dark gray and light gray colors represent the variations of irrelevant implicit error made at horizons $h-1$ and h, respectively.

[Insert Figure 4]

In general, the MSFR between two adjacent horizons decreases as the forecast day approaches the target day. Furthermore, the proportional contribution of each source of forecast revision varies over horizons. For example, at horizons longer than 8 days before the target date, adjusting for irrelevant noise accounts for most of the MSFR values. When the forecast day is within a week of the target day, more than 90% of the MSFR values are accounted for by the incorporation of newly available information. This alteration of the main source of forecast revisions can be explained by how the multi-horizon maximum temperature forecasts for each horizon are constructed. The weighting structure reported in Table 2 shows that, at horizons longer than 7 days out, forecasts are a combination of statistical and climatological forecasts derived from long-term and large-scale mathematical models. The large amount of variation in the combination forecasts that is irrelevant to the target reflects the inaccuracy of these sourcing forecasts for local maximum daily temperature when made more than 7 days out from the target day. For horizons within a week of the target, BOM's official forecasts contribute 50% of the combination forecasts. That information adoption becomes the sole source of short-term forecast revisions indicates BOM's very high short-term prediction skill for maximum daily temperature.

An understanding of the forecast revision structure across horizons can potentially help

decision makers to choose an optimal horizon forecast for their planning decisions. As analyzed above, the daily maximum temperature forecast revisions made in the second week before the target day are associated with minimal information content and hence are less likely to result in significant economic benefits for the forecast users. If early planning is preferred, they may be better off using 12-day-out revised forecasts rather than 9-day-out revised forecasts. By waiting a few days longer and bearing the opportunity costs of delayed planning, decision makers are provided with revised weather forecasts that incorporate a substantially larger amount of information. Since the information content reaches its highest level in the 6-day-out revised forecasts and then decreases as the forecast horizon shrinks, the value added using 6-day-out forecasts may be greater than that using forecasts at horizons shorter than 6 days.

5.3 Subsample evaluations

5.3.1 Effects of the NWP upgrade

The difference in forecast variabilities before and after May 22, 2012, is evident in Figure 3. On May 22, 2012, the NCEP's operational system, including the long-range numerical weather prediction models that serve as an input for the maximum temperature forecasts, were upgraded. In this section, we examine the effect of the upgrade on the multi-horizon forecast and revision structure.

We estimate the rational and implicit model using the subsample of pre-May 22, 2012, and the subsample starting on May 22, 2012. We then calculate the estimated $\text{MSFR}_{t|h-1,h}$ and its components in the two subsamples over the forecast horizons and present the results in Figure 5. The top panel is for the first subsample before the NWP model upgrade, and the bottom panel is for the second subsample after the upgrade.

[Insert Figure 5]

The sizes of $\text{MSFR}_{t|h-1,h}$ at horizons longer than 7 days after the NWP model upgrade are

approximately four times greater than those prior to the upgrade. However, the upgrade does not have much impact on the $\text{MSFR}_{t|h-1,h}$ at short horizons within 5 days before the target day is reached. The increase in $\text{MSFR}_{t|h-1,h}$ for long horizons is associated with higher $\sigma^2_{\omega_{h-1}}$ and higher $\sigma^2_{\zeta_h}$ and $\sigma^2_{\zeta_{h-1}}$, suggesting that, following the NWP model upgrades, the forecast revisions contain more relevant information but also more irrelevant noise. The upgrade does not affect the composition of the forecast revisions across horizons, and information adoption becomes the dominant source of revisions for forecasts made within 7 days of the target.

More interestingly, this NWP model upgrade advances the arrival of maximum information content. Before the upgrade, the $\text{MSFR}_{t|h-1,h}$ peaks at 6 days out, when new information adoption between two adjacent updating points is also the greatest. After the upgrade, the highest marginal information adoption occurs one day earlier, at 7 days out, and the $\text{MSFR}_{t|h-1,h}$ again exhibits a declining pattern as forecasts are made at horizons closer to the target. The NWP model upgrade provides users of maximum daily temperature forecasts with an optimal forecast horizon one day earlier than the optimal horizon prior to the upgrade. Given the same amount of information adoption in forecast revisions, since both optimal horizon forecasts contain a similar amount of newly available information, incorporating longer-horizon forecasts in planning may result in higher profits.

5.3.2 Effects of seasons

The geographical location of Melbourne leads to wide variation in the maximum temperature during warm months, making it difficult to forecast. In this section, we study whether the composition of forecast revisions across horizons is consistent over seasons. We estimate a rational and implicit model allowing for different $\sigma_{\omega_{h-1}}$ and σ_{ζ_h} between warm months and cool months. Figure 6 illustrates the values and the compositions of $\text{MSFR}_{t|h-1,h}$ over the forecast horizons, with the top panel for the warm months and the bottom panel for the cool months.

[Insert Figure 6]

The $\text{MSFR}_{t|h-1,h}$ values across all horizons in warm months are much higher than those in cool months, consistent with the fact that the forecast volatilities at all horizons are higher in warmer months. In general, the values of $\text{MSFR}_{t|h-1,h}$ decrease as the forecast horizon shrinks regardless of the season, except that the marginal forecast revisions made at 6 days out provides high $\text{MSFR}_{t|h-1,h}$.

The shares of the sources for forecast revisions are consistent between seasons. Regardless of whether the target days are in the warm or cool season, the long-horizon forecast revisions (made in the second week before the target date) are mainly due to adjustment of implicit forecast errors that are irrelevant to the target, suggesting inefficiencies in the long-horizon maximum temperature forecasts. Starting from 7-day-out forecasts, newly available information becomes the dominant attribute of the revisions, and these short-horizon maximum temperature forecasts are effective rational forecasts. Finally, the forecast horizon at which the maximum amount of marginal information arrives is the same between warm and cold months for the target.

6 Conclusion

The availability of multi-horizon forecasts of the same target offers forecast users an opportunity to investigate the revision structure across forecast horizons. This paper proposes a state space modeling approach that decomposes multi-horizon forecast errors into several unobserved components, including 1) rational forecast errors that occur due to unanticipated information related to the target; 2) implicit sources of forecast errors that are irrelevant to the target; and 3) horizon-specific bias that captures systematic under- or over-forecasts. By using this modeling approach, forecast users can explore the best fitted forecast error structure for the whole set of multi-horizon forecasts and study the key attributes of forecast revisions at

each horizon. Understanding the dynamics of forecast revision structure across horizons may help forecast users to identify the most desirable horizon forecast for their planning decisions.

Simulations demonstrate that the unobserved component modeling approach is powerful for identifying the correct forecast error structure via the use of the standard information criteria for comparing alternative error decompositions. The estimated parameters that indicate the magnitude of different sources of forecast revisions at each horizon are well behaved.

In our application, we use the proposed modeling approach to evaluate maximum daily temperature forecasts for Melbourne, Australia. Using the forecasts up to 14 days before the target, which range from February 1, 2009 to December 31, 2014, we find that these multi-horizon weather forecasts contain both rational forecast error and implicit forecast error. The variance of each type of forecast error changes along the forecast horizon, with the short-horizon revised forecasts (made up to 7 days before the target) containing more information and less irrelevant noise than the long-horizon forecasts (made in the second week before the target).

An important application of our modeling approach is to analyze the sources of forecast revisions and how the composition of these sources changes along the forecast horizon. By decomposing the value of the mean squared temperature forecast revisions between two adjacent updating points into a rational component due to adopting newly available information and an implicit component due to adjusting irrelevant noise, we show that marginal information adoption accounts for a similarly small proportion of forecast revisions made from 13 days to 9 days before the target. Information adoption becomes (effectively) the single attribute of the forecast revisions as the forecast horizon shrinks to within 7 days.

Our unobserved component modeling approach provides a means to select the best decision-making point. Early planning is beneficial but generally suffers from little relevant information being incorporated into revised forecasts made at long horizons. Our results of the forecast revision structure of the maximum daily temperature show that the forecast horizon of 6

days, when the highest amount of information is found in the revisions (using the full-sample period), could be the ideal horizon for decision making. We also find that the upgrade of the NWP models moves the occurrence of the highest marginal information adoption to the 7-day horizon and hence shifts the best decision-making time one day earlier. This improvement shows the value of investment to upgrade weather prediction systems since early planning of economic activities related to future weather conditions, including crop irrigation and harvesting and energy supply, results in potential profits for the decision makers.

References

Chang, C. L., de Bruijin, B., Franses, P. H., & McAleer, M. (2013). Analyzing fixed-event forecast revisions. *International Journal of Forecasting*, *29*, 622–627.

Christoffersen, P. F., & Diebold, F. X. (1997). Optimal prediction under asymmetric loss. *Econometric Theory*, *13*, 808–817.

Clements, M. P. (1997). Evaluating the rationality of fixed-event forecasts. *Journal of Forecasting*, *16*, 225–239.

Clements, M. P., & Taylor, N. (2001). Robust evaluation of fixed-event forecast rationality. *Journal of Forecasting*, *20*, 285–295.

Davies, A., & Lahiri, K. (1995). A new framework for analyzing survey forecasts using three-dimensional panel data. *Journal of Econometrics*, *68*, 205–227.

Davies, A., & Lahiri, K. (1999). Re-examining the rational expectations hypothesis using panel data on multi-period forecasts. In C. Hsiao, K. Lahiri, H.-F. Lee, & H. M. Pesaran (Eds.), *Analysis of panels and limited dependent variable models* (pp. 226–254). Cambridge University Press.

Davies, A., Lahiri, K., & Sheng, X. (2011). Analyzing three-dimensional panel data of forecasts. In M. P. Clements, & D. F. Hendry (Eds.), *The Oxford handbook of economic forecasting* (pp. 473–495). Oxford University Press.

Dell, M., Jones, B. F., & Olken, B. A. (2014). What do we learn from the weather? The new climate–economy literature. *Journal of Economic Literature*, *52*, 740–798.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

Isiklar, G., & Lahiri, K. (2007). How far ahead can we forecast? evidence from cross-country surveys. *International Journal of Forecasting*, *23*, 167–187.

Jacobs, J. P., & van Norden, S. (2007). Appendix to modelling data revisions: Measurement error and dynamics of true values. Mimeo, University of Groningen.

Jacobs, J. P., & Van Norden, S. (2011). Modeling data revisions: Measurement error and dynamics of true values. *Journal of Econometrics*, *161*, 101–109.

Lim, T. (2001). Rationality and analysts' forecast bias. *Journal of Finance*, *56*, 369–385.

Lovell, M. C. (1986). Tests of the rational expectations hypothesis. *The American Economic Review*, *76*, 110–124.

Mills, E. S. (1957). The theory of inventory decisions. *Econometrica*, *25*, 222–238.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica*, *29*, 315–335.

Muth, J. F. (1985). Properties of some short-run business forecasts. *Eastern Economic Journal*, *11*, 200–210.

Nordhaus, W. D. (1987). Forecasting efficiency: concepts and applications. *The Review of Economics and Statistics*, *69*, 667–674.

Patton, A., & Timmermann, A. (2007). Properties of optimal forecasts under asymmetric loss and nonlinearity. *Journal of Econometrics*, *140*, 884–918.

Patton, A. J., & Timmermann, A. (2012). Forecast rationality tests based on multi-horizon bounds. *Journal of Business & Economic Statistics*, *30*, 1–17.

Stern, H. (2007). Improving forecasts with mechanically combined predictions. *Bulletin of the American Meteorological Society*, *88*, 850–851.

Stern, H. (2008). The accuracy of weather forecasts for Melbourne, Australia. *Meteorological Applications*, *15*, 65–71.

Stern, H., & Davidson, N. E. (2015). Trends in the skill of weather prediction at lead times of 1–14 days. *Quarterly Journal of the Royal Meteorological Society*, *141*, 2726–2736.

Wang, D., & Cai, X. (2009). Irrigation scheduling: Role of weather forecasting and farmers' behavior. *Journal of Water Resource Planning and Management*, *135*, 364–373.

Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* volume 100. Academic Press.

Table 1: Simulation results of estimating the rational and implicit error component model over 1000 replications

	σ_{ω_3}	σ_{ω_2}	σ_{ω_1}	σ_{ω_0}	σ_{ζ_4}	σ_{ζ_3}	σ_{ζ_2}	σ_{ζ_1}	AIC	BIC
DGP 1	1.67	2.23	2.98	3.97	6	4	2	0		
Est. Para.	1.56	2.15	2.88	3.92	6.02	5.00	4.01	0.44	99.7%	99.7%
DGP 2	1.67	2.23	2.98	3.97	3	3	3	3		
Est. Para.	1.58	2.15	2.93	3.95	3.04	3.02	3.01	3.01	98.3%	98.3%

Notes: This table reports the averages of estimated parameters in the rational plus implicit error model over 1000 simulations. In both top and bottom panels, the first row lists the true values of the parameters used for generating rational plus implicit forecasts (that contain both rational and implicit errors) at four horizons, and the second row lists the averages of their estimated values over 1000 replications. The last two columns show the frequencies of which AIC and BIC choose the correct forecast error structure out of the alternative models. The other candidate models include the most general model that contains all three types of forecast errors and the pure rational error model. The sample size is 300.

Table 2: Forecast combination weightings for the multi-horizon daily maximum temperature forecasts in Melbourne, Australia

Horizon h	Official	Previous	Statistical	Climatology
1 to 7 days	0.50	0.25	0.25	
8 to 13 days		0.25	0.50	0.25
14 days			0.50	0.50

Notes: This table lists the weightings of the multi-horizon real-time forecasts of the daily maximum temperature of Melbourne. The weightings are provided by Stern & Davidson (2015), which explains that the forecast combination is based on official forecasts from the Australian Bureau of Meteorology, previous day's forecasts, statistical forecasts, and climatological forecasts.

Table 3: Descriptive statistics of daily maximum temperature observations and multi-horizon forecasts for Melbourne, Australia

	Full sample		Pre-NWP upgrade		Post-NWP upgrade		Warm months		Cool months	
Horizon h	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
14	20.53	4.64	20.79	4.51	20.19	4.79	24.07	3.11	17.13	3.06
13	20.57	4.72	20.79	4.50	20.30	4.98	24.14	3.29	17.17	3.08
12	20.60	4.79	20.78	4.52	20.37	5.10	24.21	3.37	17.15	3.09
11	20.61	4.82	20.78	4.53	20.40	5.17	24.24	3.42	17.14	3.12
10	20.62	4.84	20.81	4.54	20.38	5.18	24.27	3.41	17.13	3.10
9	20.60	4.79	20.83	4.55	20.31	5.07	24.20	3.40	17.15	3.09
8	20.62	4.77	20.82	4.54	20.36	5.05	24.18	3.40	17.21	3.13
7	20.74	4.95	20.82	4.59	20.65	5.37	24.31	3.78	17.33	3.23
6	20.88	5.30	21.04	5.06	20.68	5.58	24.53	4.40	17.39	3.38
5	20.93	5.45	21.11	5.24	20.70	5.69	24.57	4.71	17.45	3.47
4	20.96	5.56	21.12	5.37	20.75	5.78	24.62	4.90	17.46	3.52
3	20.97	5.62	21.11	5.44	20.79	5.85	24.62	5.04	17.48	3.56
2	20.98	5.70	21.10	5.54	20.84	5.89	24.66	5.16	17.48	3.59
1	20.99	5.76	21.13	5.62	20.81	5.93	24.67	5.28	17.47	3.60
0	21.11	6.01	21.24	5.89	20.94	6.15	24.71	5.73	17.66	3.86

Notes: This table reports the means and standard deviations of daily maximum temperature observations and their multi-horizon forecasts for Melbourne, Australia. The full sample covers 2159 target days from February 1, 2009 to December 31, 2014. The pre-NWP upgrade period spans February 1, 2009 to May 21, 2012 before the Numerical Weather Prediction models experienced a major upgrade on May 22, 2012. The post-NWP update period is from May 22, 2012 to December 31, 2014. For the southern hemisphere, we denote the days from September 21 to March 20 as warm months, and the cool months cover March 21 to September 20. The horizon h indicates the number of days before the target day that the forecasts are made. The last row, $h = 0$, represents the observed daily maximum temperature.

Table 4: Estimation results of the multi-horizon forecasts of daily maximum temperatures for Melbourne, Australia, between February 1, 2009 and December 31, 2014

		Alternative models					
		Rational		**Rational + Implicit**		**Bias +Rational + Implicit**	
Forecast Errors		Estimate	Std err	Estimate	Std err	Estimate	Std err
Rational revision	$\sigma_{\omega_{13}}$	1.568	(0.037)	0.637	(0.125)	0.667	(0.125)
	$\sigma_{\omega_{12}}$	1.654	(0.040)	0.858	(0.055)	0.857	(0.055)
	$\sigma_{\omega_{11}}$	1.564	(0.039)	0.657	(0.062)	0.657	(0.062)
	$\sigma_{\omega_{10}}$	1.534	(0.039)	0.662	(0.057)	0.662	(0.053)
	σ_{ω_9}	1.472	(0.037)	0.682	(0.053)	0.682	(0.053)
	σ_{ω_8}	1.421	(0.037)	0.992	(0.044)	0.991	(0.044)
	σ_{ω_7}	1.416	(0.036)	1.311	(0.037)	1.304	(0.037)
	σ_{ω_6}	1.560	(0.037)	1.560	(0.037)	1.554	(0.035)
	σ_{ω_5}	1.095	(0.036)	1.095	(0.036)	1.094	(0.035)
	σ_{ω_4}	0.973	(0.023)	0.973	(0.024)	0.973	(0.024)
	σ_{ω_3}	0.820	(0.022)	0.820	(0.022)	0.820	(0.022)
	σ_{ω_2}	0.728	(0.020)	0.728	(0.020)	0.727	(0.020)
	σ_{ω_1}	0.697	(0.019)	0.697	(0.019)	0.697	(0.019)
	σ_{ω_0}	1.690	(0.032)	1.691	(0.032)	1.687	(0.032)
Implicit error	$\sigma_{\zeta_{14}}$			0.994	(0.078)	0.973	(0.085)
	$\sigma_{\zeta_{13}}$			1.036	(0.038)	1.037	(0.038)
	$\sigma_{\zeta_{12}}$			1.038	(0.037)	1.038	(0.037)
	$\sigma_{\zeta_{11}}$			1.025	(0.036)	1.025	(0.036)
	$\sigma_{\zeta_{10}}$			1.010	(0.036)	1.010	(0.036)
	σ_{ζ_9}			0.874	(0.037)	0.874	(0.037)
	σ_{ζ_8}			0.534	(0.059)	0.536	(0.058)
	σ_{ζ_7}			0.005	(0.002)	0.028	(0.012)
	σ_{ζ_6}			0.003	(0.001)	0.017	(0.006)
	σ_{ζ_5}			0.002	(0.001)	0.014	(0.003)
	σ_{ζ_4}			0.003	(0.001)	0.015	(0.005)
	σ_{ζ_3}			0.002	(0.001)	0.010	(0.003)
	σ_{ζ_2}			0.002	(0.001)	0.010	(0.002)
	σ_{ζ_1}			0.003	(0.001)	0.017	(0.005)
Bias	β_{14}					-0.581	(0.066)
(smoothed states)	β_{13}					-0.531	(0.066)
	β_{12}					-0.506	(0.066)
	β_{11}					-0.494	(0.065)
	β_{10}					-0.487	(0.064)
	β_9					-0.509	(0.063)
	β_8					-0.489	(0.060)
	β_7					-0.361	(0.056)
	β_6					-0.226	(0.050)
	β_5					-0.176	(0.046)
	β_4					-0.148	(0.043)
	β_3					-0.138	(0.040)
	β_2					-0.121	(0.037)
	β_1					-0.118	(0.035)
Log likelihood		-55,914		-54,933		-55,052	
Akaike Info Criterion		111,857		109,923		110,162	
Baysian Info Criterion		111,942		110,088		110,326	

Notes: The sample covers 2159 target days from February 1, 2009 to December 31, 2014. For the daily maximum temperature of each target date, meteorological forecasts are made at a daily frequency at horizons of 14 days to 1 day before the target date. The standard errors of the estimates are in parentheses.

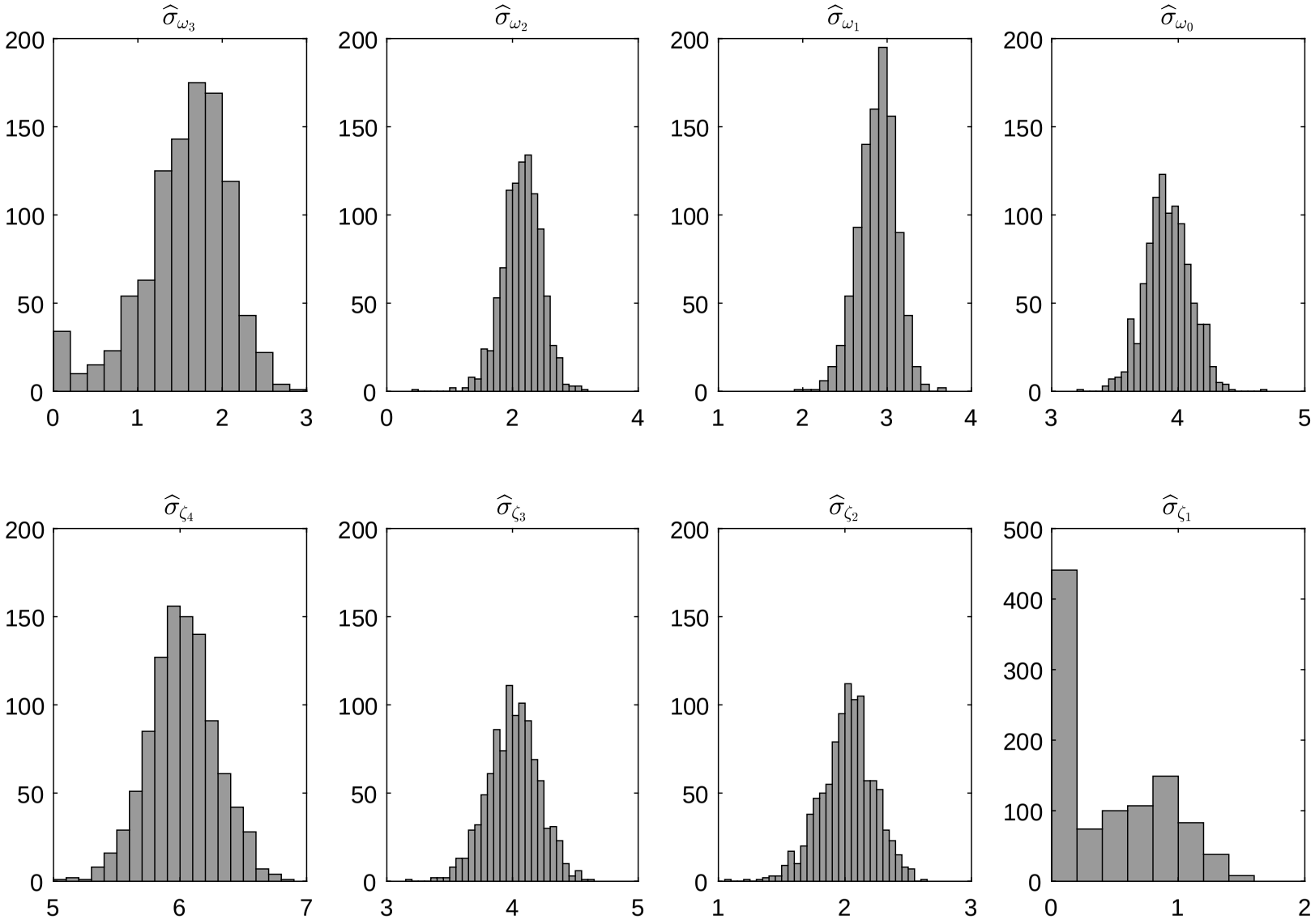

Figure 1: Simulation results of the estimated parameters in the rational plus implicit model for DGP1. The top panel shows histograms of estimated values of $\sigma_{\omega_{h-1}}$, and the bottom panel shows histograms of estimated values of σ_{ζ_h}, where $h = 4, 3, 2,$ and 1. The associated DGP has the values of σ_{ζ_h} decreasing as the forecast horizon h shrinks, specifically $\sigma_{\zeta_h} = \frac{h-1}{3}\sigma_y$. The other parameter values in DGP are discussed in the text. The sample size is 300, and the number of simulations is 1000.

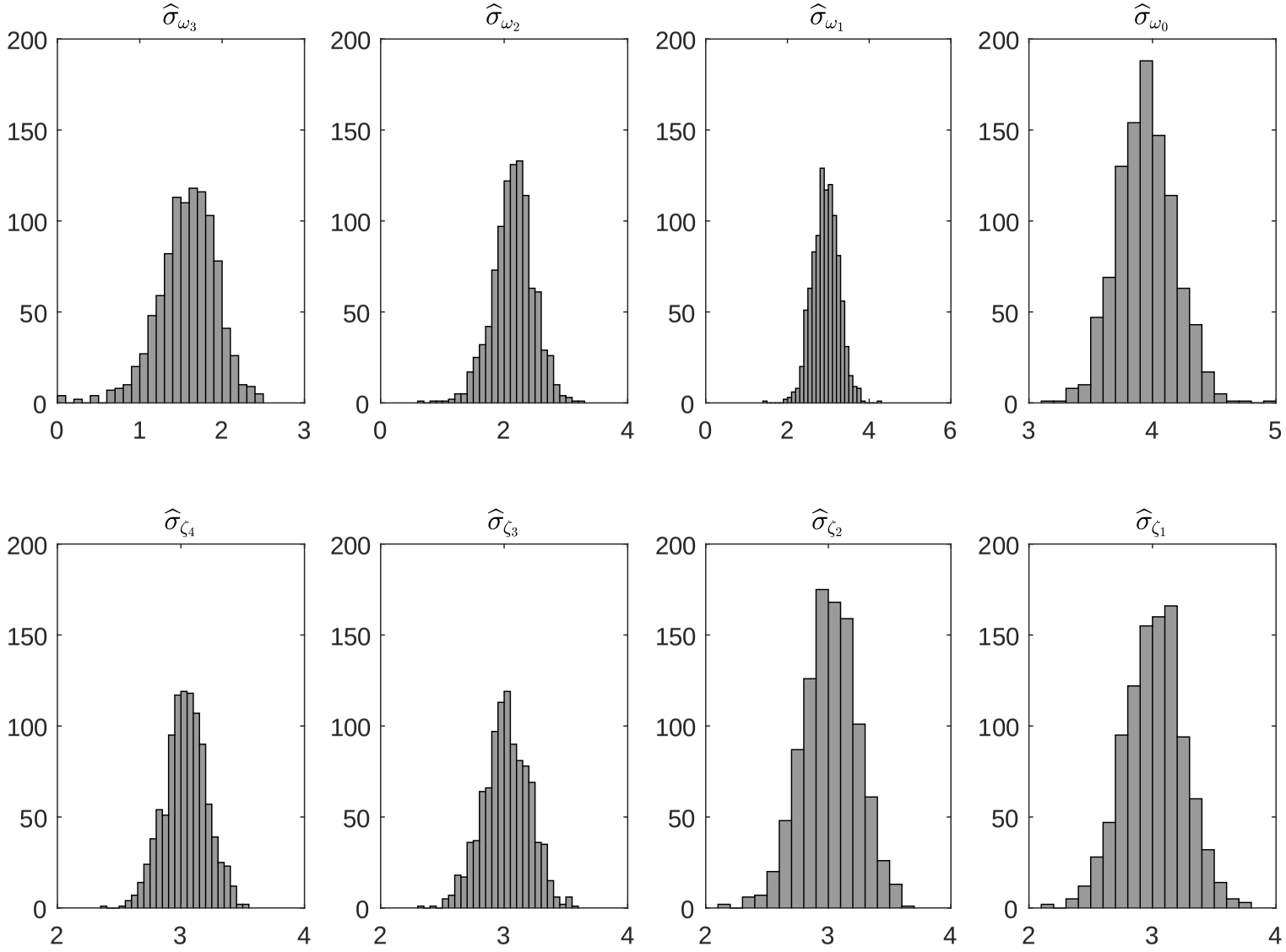

Figure 2: Simulation results of the estimated parameters in the rational plus implicit model for DGP2. The top panel shows histograms of estimated values of $\sigma_{\omega_{h-1}}$, and the bottom panel shows histograms of estimated values of σ_{ζ_h}, where $h = 4, 3, 2,$ and 1. The associated DGP has the same values of $\sigma_{\zeta_h} = 3$ over forecast horizons. The other parameter values in DGP are discussed in the text. The sample size is 300, and the number of simulations is 1000.

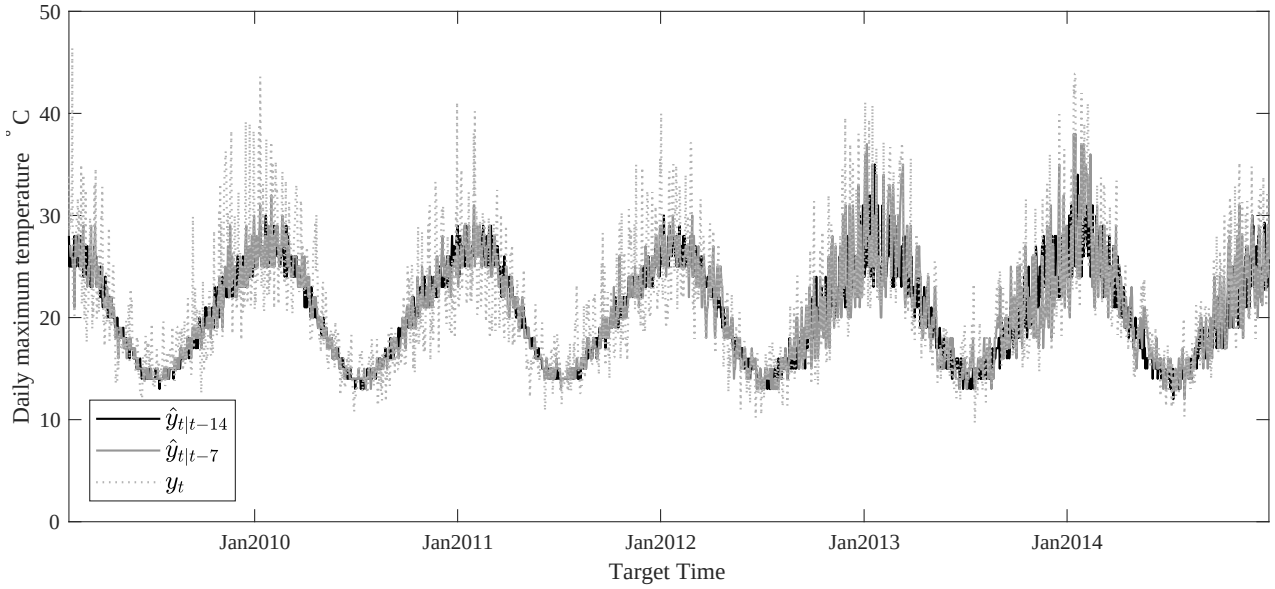

Figure 3: Daily maximum temperature observations and multi-horizon forecasts for Melbourne, Australia from February 1, 2009 to December 31, 2014

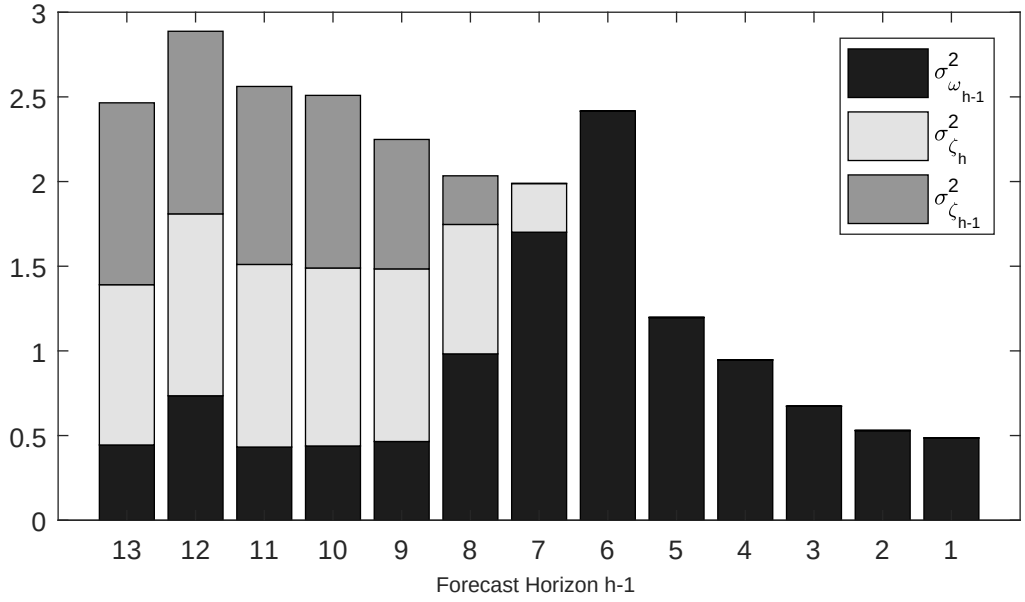

Figure 4: Estimated $\text{MSFR}_{t|h-1,h}$ and the components for multi-horizon daily maximum temperature forecasts for Melbourne, Australia, from February 1, 2009 to December 31, 2014

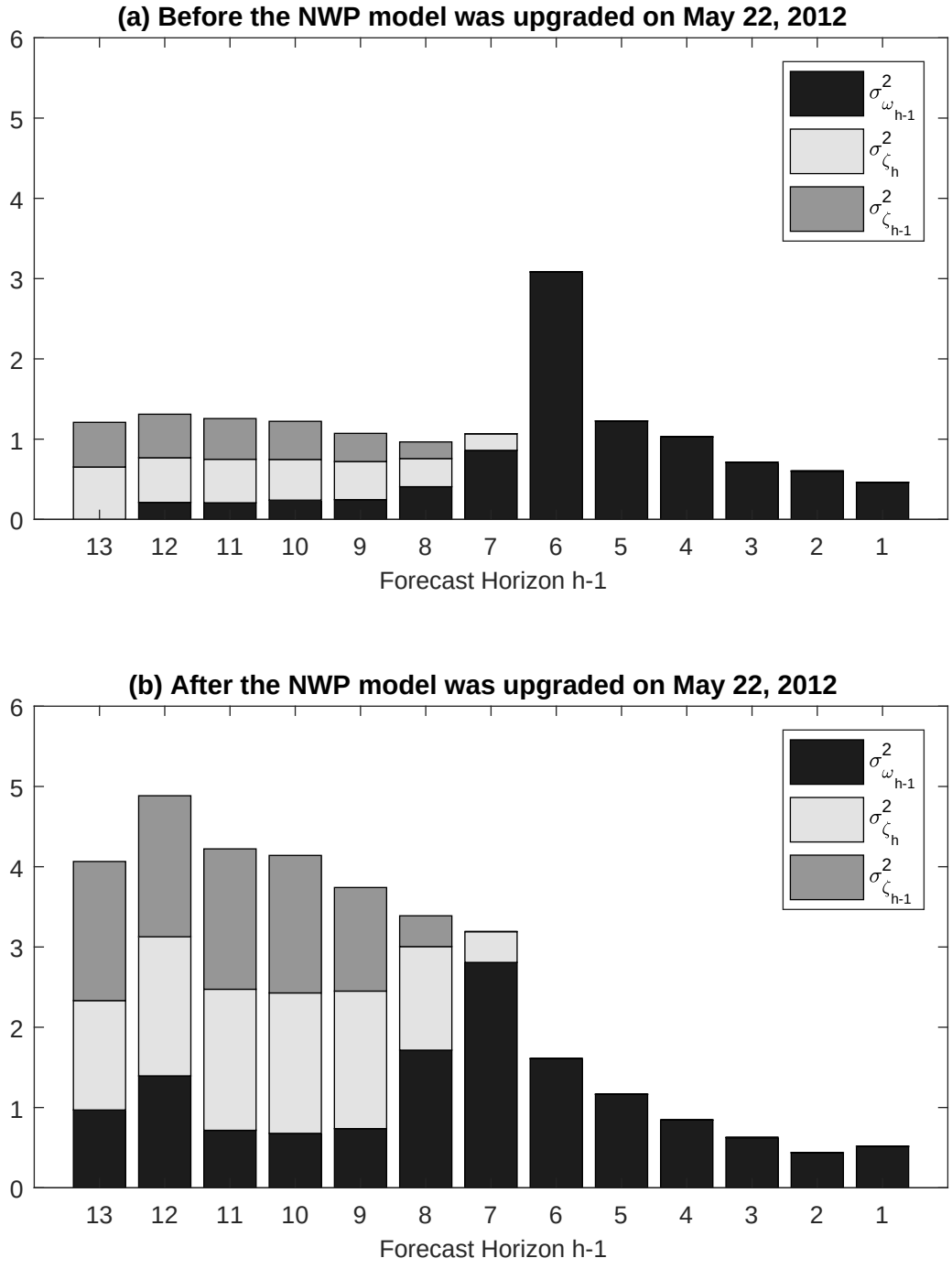

Figure 5: Comparison of estimated $\text{MSFR}_{t|h-1,h}$ and the components between the pre-NWP model upgrade and the post NWP model upgrade periods

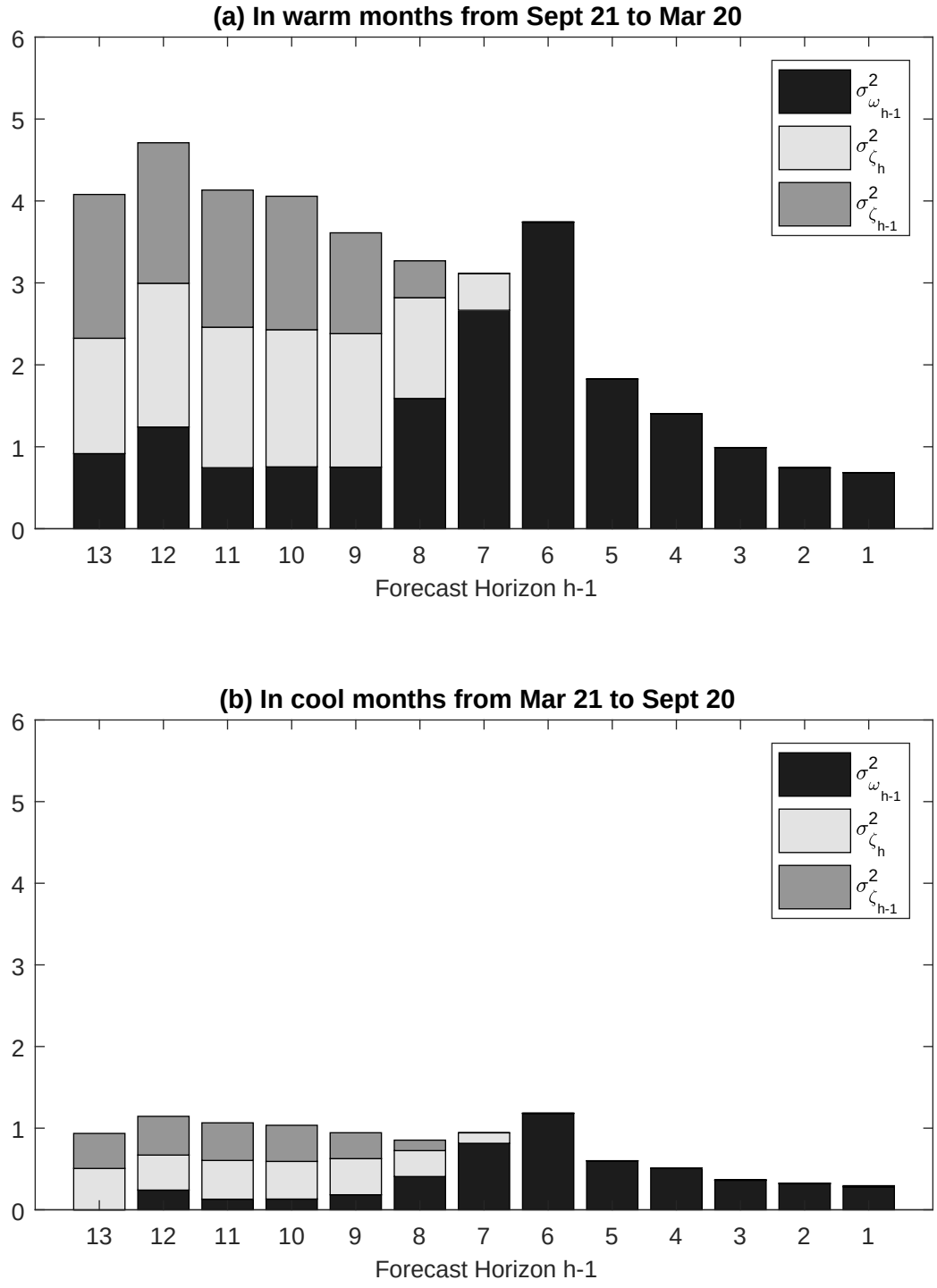

Figure 6: Comparison of estimated $\text{MSFR}_{t|h-1,h}$ and the components between warm and cool months